The Spine Journal 19 (2019) 1698–1705

Clinical Study

Which NDI domains best predict change in physical function in patients undergoing cervical spine surgery?

Michael E. Steinhaus, MD[a], Sravisht Iyer, MD[a,*], Francis Lovecchio, MD[a], Daniel Stein, BS[a], Thomas Ross, RN[a], Jingyan Yang, PhD[a,b], Virginie Lafage, PhD[a], Todd J. Albert, MD[a], Han Jo Kim, MD[a]

[a] *Department of Orthopedic Surgery, Spine Service, Hospital for Special Surgery, New York, NY, USA*
[b] *Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA*
Received 5 February 2019; revised 7 May 2019; accepted 5 June 2019

Abstract

BACKGROUND CONTEXT: Physical function is a critical aspect of patient outcomes. NDI is a widely validated outcome measure in cervical spine disease, yet to what extent its individual domains predict changes in physical function remains unknown.
PURPOSE: To examine the impact of individual NDI domains on changes in physical function in patients undergoing cervical spine surgery.
STUDY DESIGN/SETTING: Prospective Cohort Study.
PATIENT SAMPLE: Adult patients undergoing cervical spine surgery, excluding those undergoing surgery for instability due to trauma.
OUTCOME MEASURES: Absolute change in outcome measures (Patient Reported Outcomes Measurement Information System [PROMIS] Physical Function [PF], Short Form 36 [SF-36] Physical Component Score [PCS], and Neck Disability Index [NDI]) from pre- to postoperatively, correlation of NDI individual domains with PROMIS PF and SF-36 PCS (preoperatively, postoperatively, and change from pre- to postoperatively).
METHODS: Patients undergoing cervical spine surgery between 2016 and 2018 were prospectively enrolled. Patients completed questionnaires (NDI, SF-36 PCS, and PROMIS PF) preoperatively and at 6 months postoperatively. Patient demographics, including age, body mass index, Charlson Comorbidity Index, and underlying diagnoses were collected. Comparisons between NDI scores preoperatively versus postoperatively were conducted using Wilcoxon signed rank sum test. Correlations of NDI individual items and PROMIS/SF-36 were assessed using Pearson correlation. A stepwise linear regression analysis was performed to identify NDI items that are independently predictive of PROMIS PF and SF-36 PCS.
RESULTS: A total of 137 patients were included in the study, with mean age 56.9 years (range 24.4–84.9). Each of the NDI domains as well as PROMIS PF and SF-36 PCS demonstrated significant improvement following cervical spine surgery ($p<.001$). Changes in all NDI domains

FDA device/drug status: Not applicable.

Author disclosures: ***MES:*** Grant: CSRS (C). ***SI:*** Grant: CSRS (C); Grants: NASS (D). ***FL:*** Grant: CSRS (C). ***DS:*** Grant: CSRS (C). ***TR:*** Grant: CSRS (C). ***JY:*** Grant: CSRS (C). ***VL:*** Grant: CSRS (C); Stock Ownership: Nemaris Inc (20%); Consulting: Globus (D); Speaking and/or Teaching Arrangements: AO (B), J&J Depuy Synthes (C), K2M (B), Medtronic (B), Kaiser Permanete (B); Board of Directors: Nemaris Inc (20%); Research Support (Investigator Salary, Staff/Materials)^: SRS (C), HSS (C). ***TJA:*** Grant: CSRS (C); Royalties: Zimmer Biomet (F), DePuy Synthes (H), JP Medical Publishers (B), Saunders/Mosby-Elseveir (B), Thieme (B); Stock Ownership: Gentis (D), Vital 5, Bonovo Orthopedics Inc (D), Biomerix (D), InVivo Therapeutics (C), Spinicity (D), Crosstrees Medical (D), Paradigm Spine LLC (F), Invuity (C), ASIP (D), PMIG (D), Pioneer, Vertec (B); Consulting: Nuvasive (B), Facet Link (B); Board of Directors: Scoliosis Research Society (Past President 1), Hospital for Special Surgery (Surgeon-in- Chief, Medical Director, and Board of Trustees), Weill Cornell Medical College (Chairman, Department of Orthopaedic Surgery); Scientific Advisory Board/Other Office: Gentis (Medical Advisory Board), United Health Care Medical Advisory Board) (A), Clinical Orthopaedics and Related Research (Editorial Board Member) (A); Research Support (Investigator Salary, Staff/Materials)^: PCORI (F), SSG, Alan L. and Jacqueline B. Stuart Spine Research Center (E). ***HJK:*** Grant: CSRS (C); Royalties: Zimmer Biomet (E), K2M (B); Scientific Advisory/Other Office: AO Spine (B); Research Support (Investigator Salary, Staff/Materials): ISSGF; Fellowship Support: AO Spine.

Investigation performed at the Department of Orthopedic Surgery, Hospital for Special Surgery, New York, NY.

* Corresponding author. Department of Orthopedic Surgery, Hospital for Special Surgery, 535 East 70th St, New York, NY 10021, USA. Tel.: (212) 606-1875.

E-mail addresses: iyers@hss.edu, sravisht@gmail.com (S. Iyer).

https://doi.org/10.1016/j.spinee.2019.06.006

demonstrated significant negative correlation with changes in PROMIS PF, with recreation (R=−0.537, p<.001), work (R=−0.514, p<.001), and pain intensity (R=−0.488, p<.001) having the greatest negative correlation. Changes in all NDI domains demonstrated significant negative correlation with changes in SF-36 PCS, with recreation (R=−0.451, p<.001), work (R=−0.443, p<.001), lifting (R=−0.373, p<.001), and driving (R=−0.373, p<.001) having the greatest negative correlation. For PROMIS PF, the NDI domains that were independently associated with changes in PF were work (R=−0.092, p=.001), pain intensity (R=−0.089, p=.003), and recreation (R=−0.067, p=.004). For SF-36 PCS, the NDI items that were independently associated with changes in PCS were work (R=−0.269, p=.003) and recreation (R=−0.215, p=.002).

CONCLUSIONS: All NDI domains improve significantly after cervical spine surgery and demonstrate significant correlation with changes in PROMIS PF and SF-36 PCS. The work, recreation, and pain intensity domains were the only independent predictors of physical function changes postoperatively. Considering physical function, our findings highlight the importance of presenting changes in individual NDI domains in addition to the total score.

Keywords: Cervical spine surgery; Cervical myelopathy; Cervical radiculopathy; Patient-reported outcomes; Physical function; Outcomes reporting

Introduction

Patient-reported outcome measures (PROMs) have emerged as one of the tools of assessing quality of interventions in spine surgery [1]. One of the most common PROMs used in cervical spine disease, the Neck Disability Index (NDI) consists of 10 different domains, including pain intensity, personal care, lifting, reading, headaches, concentration, work, driving, sleeping, and recreation [2]. Each item receives a score from 0 to 5, with 0 representing the least disability and 5 representing the greatest, with an aggregate score calculated from these 10 domains.

Originally modified from the Oswestry Disability Index (ODI), the NDI was initially validated by Vernon and Mack in 1991 by assessing concurrent validity with the visual analog scale as well as with the McGill Pain Questionnaire [2]. The NDI is now widely used and validated in cervical spine pathology. It has been studied in patients treated conservatively for neck pain [3,4], those undergoing cervical fusion for degenerative disease [5], in Food & Drug Administration Investigational Device Exemption trials assessing cervical disc arthroplasty [6–9], and has been adapted for use in a multitude of languages. Although it is the primary disease-specific instrument used in these conditions, it is not without drawbacks. Certain studies have demonstrated limitations in patients with cervical pathology, including poor construct validity and limited test-retest reliability [4], floor effects [10], and multidimensionality [10,11], which can make interpreting results a challenge.

In addition to NDI, other instruments are used in assessing outcomes in the cervical spine. Rather than a disease-specific instrument, the Short Form 36 (SF-36) is a tool which measures general health outcomes to enable comparison across various conditions and interventions [12]. As a part of the assessment, SF-36 explores physical, mental, and social functioning, with the Physical Component Score (PCS) measuring the impact of physical components on overall health. In addition to demonstrating validity with respect to generic measures [13,14], it has been applied in disease-specific settings including cervical spinal pathology [5,15–18]. The SF-36 PCS specifically has been studied in patients with mild degenerative myelopathy [17], those undergoing fusion for degenerative disease [5], and those undergoing disc arthroplasty [18].

Another measure of general health, the Patient Reported Outcomes Measurement Information System (PROMIS), was developed more recently by the National Institutes of Health to address shortcomings associated with existing PROMs [19]. Initially validated in general population and clinical groups [19], PROMIS has been compared across chronic conditions [20], and since has been validated in many disease-specific settings, such as in those starting treatment for depression [21], patients starting disease-modifying treatment in rheumatoid arthritis [22], as well as in patients with spinal pathology [23]. Unlike these other instruments, certain versions of PROMIS use computer adaptive testing (CAT) which allows the use of fewer questions, reducing burden to the patient and, as a general health measure, adds the benefit of enabling comparison across different pathologies and interventions. Multiple studies have tested PROMIS in cervical surgery, finding that it may represent an improvement over legacy instruments with increased efficiency and responsiveness [24–27].

Although several instruments are available for use in cervical spine patients, in the current resource- and time-constrained environment not all patients and clinicians complete all of these measures. The NDI is widely used, yet the majority of studies simply report a composite score, obscuring the results of individual domains. As physical function is a critical aspect of patient outcome that the SF-36 PCS and PROMIS PF attempt to address, the objective of this study was to determine which NDI domains correlate with changes in physical function. Additionally, we

aimed to determine which of these domains were independent predictors of such changes. Demonstrating which items are predictive of such changes will assist surgeons and patients in understanding physical function prognosis for those undergoing surgery for cervical spinal disease.

Material and methods

Before initiation, the Institutional Review Board approved the study and informed consent was obtained. All adult (≥18 years) patients undergoing cervical spine surgery at a single institution were screened for enrollment. Patients were included if they were undergoing surgery for cervical degenerative disease or deformity by one of seven spine surgeons between 2016 and 2018. Patients were excluded if they were not fluent English speakers or were undergoing surgery for cervical instability due to trauma.

Questionnaires

Patients completed NDI, SF-36 PCS, and PROMIS Physical Function (PF) CAT questionnaires pre- and postoperatively. The PROMIS Bank v1.2 - Physical Function CAT was utilized. Preoperatively, surveys were completed at the office visit at which patients were indicated for surgery or in the preoperative holding area. Postoperative assessment was performed at 6 months after surgery with a 90-day window allotted for survey completion. Six months was selected as the majority of gains in PF following cervical surgery have been shown to occur within the first 6 months [16,28,29]. All questionnaires were administered electronically via the Assessment Center (www.assessmentcenter.net). Instruments were administered in a randomized fashion to minimize questionnaire fatigue or ordering and any resulting bias.

Data collection

Patient demographics, including age, body mass index, Charlson Comorbidity Index, and underlying symptoms were collected. Patient responses to questionnaires were collected.

Patient outcome instruments

NDI

The NDI reports disability due to neck pain for a variety of activities of daily living, including 10 domains (sleeping, reading, driving, etc.). Higher NDI scores indicate greater disability (ie, worse outcome).

SF-36 PCS

The SF-36 is a measure of general health, with the PCS measuring the impact of physical components on overall health. Higher SF-36 scores indicate better functional status.

PROMIS PF CAT

The PROMIS PF CAT measures patient "self-reported capability rather than the actual performance of physical activities" [30]. These include functioning of upper/lower extremities, neck, and back, and the performance of activities of daily living. The computer adaptive test uses 4–12 questions, using patient responses to determine subsequent questions. Questions continue until the standard error falls below a specified value (eg, <3.0 t-score) or the patient has answered a maximum of 12 questions [31]. The PF CAT draws from an item bank consisting of 121 total questions. Higher PF scores indicate better functional status.

Statistical analysis

Descriptive statistics were summarized as mean (SD) for normally distributed continuous variables, median (IQR) for non-normally distributed continuous variables and count (frequency) for categorical variables. Comparisons between NDI scores preoperatively versus postoperatively were conducted using Wilcoxon signed rank sum test. Correlations of NDI individual items and PROMIS/SF-36 were assessed using Pearson correlation. In order to identify NDI items that are independently predictive of PROMIS PF and SF-36 PCS, a stepwise linear regression analysis was performed. The significance level was set at $p<.05$. For analyses, the NDI raw scores were multiplied by two to create scores out of 100. All the analyses were conducted using Stata SE 14.0 (StataCorp, College Station, TX, USA).

Theory

The majority of studies simply report NDI as a composite score, with the results of the individual domains not being detailed. Although this reporting is consistent with the intended use of the NDI, it is useful to understand the underlying components of that score, as each item captures a different aspect of function. Whether and to what extent the individual domains of NDI predict changes in PF is critical in understanding the effects of different interventions in cervical spine surgery. Identifying these domains is useful to patients and clinicians alike in providing prognostic information about expected changes in PF after intervention.

Results

Demographics and outcomes scores

A total of 164 patients completed preoperative surveys were available for 6-month follow-up; 137 of the 164 patients (83.5%) had 6-month follow-up data available for analysis (Table 1). There were 82 male (59.8%) and 55 female (40.2%) patients with a mean age of 56.9 years (range 24.4–84.9). Mean body mass index was 27.4 kg/m^2 (range 17.5–43.9) and mean Charlson Comorbidity Index was 1.7 (range 0–6). The most common underlying

Table 1
Patient demographics

	Mean (range)	SD
Age, yrs	56.9 (24.4–84.9)	13.1
	N	%
Male	82	59.8%
Female	55	40.2%
BMI	27.4 (17.5–43.9)	4.8
BMI, %		
Underweight/ Normal	45	32.8%
Overweight	53	38.7%
Obese	39	28.5%
CCI	1.7 (0–6)	1.5
Diagnosis	N	%
Myelopathy	23	16.8%
Radiculopathy	32	23.4%
Myeloradiculopathy	29	21.2%
Pseudarthrosis	2	1.5%
OPLL	3	2.2%
Facet arthropathy	25	18.2%
ASD	5	3.7%
Other	18	13.1%

SD, standard deviation; BMI, body mass index; CCI, Charlson Comorbidity Index.

symptom was radiculopathy (n=32 patients; 23.4%), followed by myeloradiculopathy (n=29; 21.2%) and myelopathy (n=23; 16.8%). Patients who did not complete 6 month surveys and were therefore not included did not differ in any baseline characteristic from the included group.

Each of the NDI domains demonstrated significant improvement following cervical spine surgery (p<.001; Table 2). The recreation (Δ=25.3), sleeping (Δ=23.3), and pain intensity (Δ=20.6) domains demonstrated the greatest absolute improvement. Significant improvements were also noted in PROMIS PF and SF-36 PCS (improvements of 7.0 and 16.9, respectively; p<.001). We found that 59% of patients met the calculated minimal clinically important difference (MCID) of 4.5 for PROMIS PF.

Correlations between NDI and PROMIS PF

Preoperatively, all NDI domains demonstrated significant negative correlation (ie, the greater the NDI, the lower the PROMIS PF) with preoperative PROMIS PF scores (Table 3). Of the domains, work (R=−0.622, p<.001), personal care (R=−0.597, p<.001), and lifting (R=−0.551, p<.001) showed the greatest negative correlation with preoperative PROMIS PF, whereas headaches (R=−0.261, p=.002), pain intensity (R=−0.369, p<.001), and reading (R=−0.408, p<.001) showed the least.

Postoperatively, all NDI domains showed significant negative correlation with postoperative PROMIS PF scores (p<.001 for all) (Table 3). Of the domains, work (R=−0.677), recreation (R=−0.644), and lifting (R=−0.568) showed the greatest negative correlation, whereas headaches (R=−0.425), sleeping (R=−0.427), and driving (R=−0.440) showed the least.

Changes in all NDI domains demonstrated significant negative correlation with changes in PROMIS PF (Table 3). Of the domains, recreation (R=−0.537, p<.001), work (R=−0.514, p<.001), and pain intensity (R=−0.488, p<.001) showed the greatest negative correlation, whereas headaches (R=−0.187, p<.033), personal care (R=−0.300, p=.001), and sleeping (R=−0.360, p<.001) showed the least.

Correlations between NDI and SF-36 PCS

Preoperatively, eight of 10 NDI domains demonstrated significant negative correlation with preoperative SF-36 PCS scores (Table 4). Of the domains, personal care (R=−0.481, p<.001), work (R=−0.466, p<.001), and lifting (R=−0.427, p<.001) showed the greatest negative correlation with preoperative SF-36 PCS, whereas reading (R=0.245, p=.004), pain intensity (R=−0.121, p=.162), and headaches (R=−0.159, p<.066) showed the least. Neither pain intensity nor headaches showed a statistically

Table 2
Outcomes scores.

Item	Median (IQR) (Preoperative)	Median (IQR) (Postoperative)	Mean difference	p Value*
Pain intensity	40 (20, 60)	10 (0, 20)	20.6	<.001
Personal care	20 (0, 20)	0 (0, 0)	9.6	<.001
Lifting	40 (20, 60)	20 (0, 60)	13.1	<.001
Reading	40 (20, 60)	20 (0, 20)	15.6	<.001
Headaches	20 (0, 40)	20 (0, 20)	10.2	<.001
Concentration	20 (0, 40)	0 (0, 20)	12.9	<.001
Work	40 (20, 60)	0 (0, 30)	19.8	<.001
Driving	40 (20, 60)	0 (0, 20)	18.5	<.001
Sleeping	40 (20, 60)	20 (0, 40)	23.3	<.001
Recreation	40 (20, 80)	20 (0, 40)	25.3	<.001
PROMIS PF	41.3 (36.5, 46.1)	48.6 (43.3, 54.3)	−7.0	<.001
SF-36 PCS	49.8 (39.3, 65)	78.3 (54.3, 91.0)	−16.9	<.001

IQR, interquartile range; PROMIS, Patient Reported Outcomes Measurement Information System; SF-36, Short Form 36; PF, physical function; PCS, Physical Component Score.

Table 3
Correlation of NDI items with PROMIS PF

Item	Preoperative		Postoperative		Difference	
	Correlation Coefficient	p Value	Correlation Coefficient	p Value	Correlation Coefficient	p Value
Pain intensity	−0.369	<.001	−0.519	<.001	−0.488	<.001
Personal care	−0.597	<.001	−0.566	<.001	−0.3	.001
Lifting	−0.551	<.001	−0.568	<.001	−0.395	<.001
Reading	−0.408	<.001	−0.485	<.001	−0.389	<.001
Headaches	−0.261	.002	−0.425	<.001	−0.187	.033
Concentration	−0.448	<.001	−0.469	<.001	−0.361	<.001
Work	−0.622	<.001	−0.677	<.001	−0.514	<.001
Driving	−0.477	<.001	−0.44	<.001	−0.395	<.001
Sleeping	−0.476	<.001	−0.427	<.001	−0.36	<.001
Recreation	−0.541	<.001	−0.644	<.001	−0.537	<.001

NDI, Neck Disability Index; PROMIS, Patient Reported Outcomes Measurement Information System; PF, physical function.

Table 4
Correlation of NDI items with SF-36 PCS

Item	Preoperative		Postoperative		Difference	
	Correlation Coefficient	p Value	Correlation Coefficient	p Value	Correlation Coefficient	p Value
Pain intensity	−0.121	.162	−0.411	<.001	−0.202	.02
Personal care	−0.481	<.001	−0.612	<.001	−0.354	<.001
Lifting	−0.427	<.001	−0.496	<.001	−0.373	<.001
Reading	0.245	.004	−0.478	<.001	−0.309	<.001
Headaches	−0.159	.066	−0.373	<.001	−0.174	.048
Concentration	−0.384	<.001	−0.542	<.001	−0.281	.001
Work	−0.466	<.001	−0.7	<.001	−0.443	<.001
Driving	−0.271	.002	−0.435	<.001	−0.373	<.001
Sleeping	−0.306	<.001	−0.431	<.001	−0.222	.011
Recreation	−0.307	<.001	−0.591	<.001	−0.451	<.001

NDI, Neck Disability Index; SF-36, Short Form 36; PCS, Physical Component Score.

significant correlation with preoperative SF-36 PCS scores, whereas reading demonstrated a significant *positive* correlation (ie, opposite effect).

Postoperatively, all NDI domains showed significant correlation with postoperative SF-36 PCS scores (p<.001 for all) (Table 4). Of the domains, personal care (R=−0.612), recreation (R=−0.591), and concentration (R=−0.542) showed the greatest negative correlation, whereas headaches (R=−0.373), pain intensity (R=−0.411), and sleeping (R=−0.431) showed the least.

Changes in all NDI domains demonstrated significant negative correlation with changes in SF-36 PCS (Table 4). Of the domains, recreation (R=−0.451, p<.001), work (R=−0.443, p<.001), lifting (R=-0.373, p<.001), and driving (R=−0.373, p<.001) showed the greatest negative correlation, whereas headaches (R=−0.174, p=.048), pain intensity (R=−0.202, p=.02), and sleeping (R=−0.222, p=.011) showed the least.

Predictive NDI domains

Linear regression was performed for change in NDI items versus change in PROMIS PF and SF-36 PCS (Table 5). For PROMIS PF, the NDI domains that were independently associated with changes in PF were work (R=−0.092, p=.001; confidence interval [CI]: −0.148 to −0.036), pain intensity (R=−.089, p=.003; CI: −0.147 to −0.032), and recreation (R=−0.067, p=.004; CI: −0.112 to −0.022). For SF-36 PCS, the NDI items that were independently associated with changes in PCS were work (R=−0.269, p=.003; CI: −0.444 to −0.093) and recreation (R=−0.215, p=.002; CI: −0.351 to −0.079).

Discussion

Our study has several important findings. First, each of the NDI domains demonstrated significant improvement following cervical spine surgery, with recreation, sleeping, and pain intensity domains showing the greatest absolute improvement. Similarly, significant improvements were also seen in PROMIS PF and SF-36 PCS from pre- to postoperatively. Changes in each NDI domain demonstrated correlation with changes in PROMIS PF and SF-36 PCS, with certain domains showing greater degrees of correlation than others. For example, changes in the recreation and work domains had high negative correlations with changes in physical function for both PROMIS and SF-36, whereas changes in the headaches and sleeping domains showed

Table 5
Linear regression modeling

Item	PROMIS PF Coefficient (95% CI)	p Value	SF-36 PCS Coefficient (95% CI)	p Value
Pain intensity	−0.089 (−0.147, −0.032)	.003		
Personal care	-		-	
Lifting	-		-	
Reading	-		-	
Headaches	-		-	
Concentration	-		-	
Work	−0.092 (−0.148, −0.036)	.001	−0.269 (−0.444, −0.093)	.003
Driving	-		-	
Sleeping	-		-	
Recreation	−0.067 (−0.112, −0.022)	.004	−0.215 (−0.351, −0.079)	.002

NDI, Neck Disability Index; PROMIS, Patient Reported Outcomes Measurement Information System; PF, physical function; SF-36, Short Form 36; PCS, Physical Component Score; CI, confidence interval.

lower correlations with changes in physical function for these instruments. Finally, regression analysis was performed to identify specific items that were independently associated with changes in PROMIS PF and SF-36 PCS. The work and recreation domains demonstrated an independent correlation with both of these measures of physical function, and the pain intensity domain was additionally independently correlated with changes in PROMIS-PF, suggesting that these domains are the most important contributors to change in physical function after cervical spine surgery.

These findings have important implications. Patient-reported outcomes have gained in popularity and relevance, not only providing information on expectations to patients and physicians, but also informing payors assessing the outcomes of a specific intervention. As these key stakeholders become increasingly reliant on PROMs in decision-making, ensuring that these instruments accurately reflect changes in a patient's condition becomes ever more vital. If an instrument is not valid, responsive, or reliable in a certain patient population with a specific diagnosis, changes (or lack thereof) in an instrument may not reflect the true impact of an intervention, leading patients, physicians, and payors to make misinformed management decisions. Furthermore, for outcome measures that include multiple domains, it is similarly important to evaluate each item—not only the composite score—as these domains can have heterogeneous effects on the outcome of interest. Here we show that while all domains of the NDI were significantly correlated with PROMIS-PF and SF-36 PCS, the magnitude of each correlation varied and only three of the domains (recreation, work, and pain intensity) were independently associated with physical function changes. In our resource- and time-constrained environment, many clinicians, and patients do not have the resources and/or time to complete all of the different instruments that have been validated to test different aspects of disease state. As NDI is a widely used measure in cervical spine, understanding which domains predict changes in physical function is valuable to patients and clinicians alike.

Unfortunately, data comparing NDI to physical function are scarce. An older study by Riddle and Stratford compared the construct validity and sensitivity to change of NDI and SF-36, finding that there was construct validity and sensitivity to change of NDI and the SF-36 PCS and MCS, noting overlap between the measures and that the NDI measures both mental and physical factors [32]. A more recent study by Owen et al. reported a strong negative correlation between PROMIS PF and NDI for patients undergoing surgery for cervical myelopathy; however, these authors only tested the NDI composite score and did not include details on the different domains which, as shown in this study, have heterogeneous effects [26].

Because the NDI remains the de facto gold standard for patients undergoing cervical spine surgery, it is valuable for clinicians to understand how to interpret changes in this score. Although the composite NDI score has been validated and should remain the most widely reported measure, our results demonstrate that three domains (work, recreation, and pain intensity) also predict changes in patients' physical function. While work and recreation domains remained in the model for both PROMIS PF and SF-36 PCS, pain intensity was additionally correlated with PROMIS PF but not SF-36 PCS. The reasons for this difference are unknown, although it is possible that differences in question stem wording about pain may contribute to this discrepancy. For example, the SF-36 asks about "bodily pain" which may be interpreted by patients as different from "neck" or "arm" pain, resulting in correlational differences with the NDI. Regardless of this difference, the concept that the selected domains of the NDI can predict physical function is important. In the case of studies that

look only at NDI and do not examine other measures of physical function, focusing on these two or three domains should provide information on changes in a patient's PF post-intervention [33].

Similar work has been performed for pathology of the lumbar spine, assessing how well ODI domains predict changes in PF in degenerative lumbar spondylolisthesis [33]. In their work, Murphy et al. report that the greatest improvement after surgery was found in the standing, sex life, and social life domains, but that after regression analysis the lifting, standing, and traveling domains remained as independent predictors of change in PF (using the SF12 PCS). They found that using the regression model with these items better predicted changes in SF12 PCS than did the composite ODI score. To confront this challenge, the authors suggest that future studies report on improvement in individual domains, as certain domains may not be relevant to the study question. Another study by Djurasovic et al. assessed changes in ODI and SF-36 PCS in patients undergoing lumbar fusion, and found that only pain intensity, walking, and social life domains were independent predictors of SF-36 PCS, giving further evidence of the importance of assessing individual items [34].

There are several limitations of our study. The patient sample is relatively homogeneous and reflects the experience at a single institution, which should be considered before generalizing the findings, and the results are not adjusted for potential confounders. Additionally, this study does not take into account any mental component, which is a critical contributor to a patient's overall function and disability. The follow-up time of 6 months could be considered a further drawback, as many recommend 2 years of follow-up for postoperative outcomes. Nevertheless, in cervical surgery it has been shown that postoperative gains in PF largely plateau within the first 6 months postoperatively for anterior [28,29] and posterior [16] surgery, although one study showed that NDI continued to improve up to 2 years in the posterior group [16]. Furthermore, patient responses can be impacted by fatigue and the order in which questionnaires are completed. To address this problem, we collected data through an assessment center, randomizing the order in which the outcome measures were administered in order to limit any chance of this type of bias. Finally, with the exception of instability due to trauma, we included all patients undergoing cervical surgery and it is possible that our results would vary by a number of factors (age, duration and types of symptoms, type of work, number of levels, procedure type, etc.). Although the correlation of NDI items with measures of physical function may have some variation across these factors, studies have shown that PROMIS and NDI demonstrate strong correlation across a variety of pathologies/procedure types [23–27,35,36]. Thus, we did not feel that our results would substantially change with these variables, and stratifying by these factors limits the size of our comparator groups, limiting our ability to perform meaningful analyses. Nevertheless, it is possible that certain procedural or symptom-based factors could impact these correlations, which could be a topic of future research in a more homogeneous or larger cohort.

One final potential limitation of the study is the deconstruction of the NDI composite into its individual domains. The NDI is widely validated as a composite score (ie, not based on its individual domains) as a measure of neck disability and not specifically for PF after cervical surgery. The results of this study are not intended to invalidate the NDI as an outcome measure. Rather, recognizing its importance and widespread use, the goal of the study was to understand which domains of the NDI correlate with and predict changes in PF, a critical aspect of patient outcome.

Conclusions

In conclusion, all domains of the NDI improve significantly after surgery for cervical spine disease, with recreation, sleeping, and pain intensity items having the greatest change. Additionally, all domains demonstrated significant correlation with changes in PROMIS PF and SF-36 PCS scores. Regression analysis identified work, recreation, and pain intensity domains as being independent predictors or changes in physical function, suggesting that changes in these domains have the greatest impact on a patients' PF postoperatively. Future studies should consider the heterogeneous effects of different NDI items, and where possible reporting the results of individual domains when evaluating impact on PF.

References

[1] Rihn JA, Currier BL, Phillips FM, Glassman SD, Albert TJ. Defining the value of spine care. J Am Acad Orthop Surg 2013;21:419–26.
[2] Vernon H, Mior S. The Neck Disability Index: A study of reliability and validity. J Manipulative Physiol Ther 1991;14:409–15.
[3] Pool JJM, Ostelo RWJG, Hoving JL, Bouter LM, De Vet HCW. Minimal clinically important change of the neck disability index and the numerical rating scale for patients with neck pain. Spine 2007;32:3047–51.
[4] Young IA, Dunning J, Butts R, Mourad F, Cleland JA. Reliability, construct validity, and responsiveness of the neck disability index and numeric pain rating scale in patients with mechanical neck pain without upper extremity symptoms. Physiother Theory Pract 2018:1–8.
[5] Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck Disability Index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. Spine J 2010;10:49–74.
[6] Sasso RC, Smucker JD, Hacker RJ, Heller JG. Clinical outcomes of BRYAN cervical disc arthroplasty: a prospective, randomized, controlled, multicenter trial with 24-month follow-up. J Spinal Disord Tech 2007;20:481–91.
[7] Heller JG, Sasso RC, Papadopoulos SM, Anderson PA, Fessler RG, Hacker RJ, et al. Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. Spine 2009;34:101–17.

[8] Janssen ME, Zigler JE, Spivak JM, Delamarter RB, Darden BV. ProDisc-C total disc replacement versus anterior cervical discectomy and fusion for single-level symptomatic cervical disc disease. J Bone Jt Surg Am 2015;97:1738–47.
[9] Davis RJ, Nunley PD, Kim KD, Hisey MS, Jackson RJ, Bae HW, et al. Two-level total disc replacement with Mobi-C cervical artificial disc versus anterior discectomy and fusion: a prospective, randomized, controlled multicenter clinical trial with 4-year follow-up results. J Neurosurg Spine 2015;22:15–25.
[10] Hung M, Cheng C, Hon SD, Franklin JD, Lawrence BD, Neese A, et al. Challenging the norm: Further psychometric investigation of the neck disability index. Spine J 2015;15:2440–5.
[11] Van Der Velde G, Beaton D, Hogg-Johnston S, Hurwitz E, Tennant A. Rasch analysis provides new insights into the measurement properties of the neck disability index. Arthritis Care Res 2009;61:544–51.
[12] Ware JEJ, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Med Care 1992;30:473–83.
[13] Brazier JE, Harper R, Jones NM, O'Cathain A, Thomas KJ, Usherwood T, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992;305:160–4.
[14] McHorney CA, Ware JE, Raczek AE. The MOS 36-item short-form health survey (Sf-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247–63.
[15] Ko S, Choi W, Chae S, Kwon J, Lee Y. Correlation between Short-form 36 scores and neck disability index in patients undergoing anterior cervical discectomy and fusion. *Asian* Spine J 2018;12:691–6.
[16] Seng C, Tow BPB, Siddiqui MA, Srivastava A, Wang L, Yew AKS, et al. Surgically treated cervical myelopathy: a functional outcome comparison study between multilevel anterior cervical decompression fusion with instrumentation and posterior laminoplasty. Spine J 2013;13:723–31.
[17] Badhiwala JH, Witiw CD, Nassiri F, Jaja BNR, Akbar MA, Mansouri A, et al. Patient phenotypes associated with outcome following surgery for mild degenerative cervical myelopathy: a principal component regression analysis. Spine J 2018;18:2220–31.
[18] Heary RF, Goldstein IM, Getto KM, Agarwal N. Solid radiographic fusion with a nonconstrained device 5 years after cervical arthroplasty. J Neurosurg Spine 2014;21:951–5.
[19] Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010;63:1179–94.
[20] Rothrock NE, Hays RD, Spritzer K, Yount SE, Riley W, Cella D. Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). J Clin Epidemiol 2010;63:1195–204.
[21] Pilkonis PA, Yu L, Dodds NE, Johnston KL, Maihoefer CC, Lawrence SM. Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS®) in a three-month observational study. J Psychiatr Res 2014;56:112–9.
[22] Wohlfahrt A, Bingham CO, Marder W, Phillips K, Bolster MB, Moreland LW, et al. Responsiveness of Patient Reported Outcomes Measurement Information System (PROMIS) measures in RA patients starting or switching a DMARD. Arthritis Care Res 2019;71:521–9.
[23] Hung M, Saltzman CL, Voss MW, Bounsanga J, Kendall R, Spiker R, et al. Responsiveness of the PROMIS, NDI and ODI instruments in patients with spinal disorders. Spine J 2019;19:34–40.
[24] Boody BS, Bhatt S, Mazmudar AS, Hsu WK, Rothrock NE, Patel AA. Validation of Patient-Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests in cervical spine surgery. J Neurosurg Spine 2018;28:268–79.
[25] Papuga MO, Mesfin A, Molinari R, Rubery PT. Correlation of PROMIS physical function and pain CAT instruments with Oswestry disability index and neck disability index in spine patients. Spine 2016;41:1153–9.
[26] Owen RJ, Zebala LP, Peters C, McAnany S. PROMIS physical function correlation with NDI and mJOA in the surgical cervical myelopathy patient population. Spine 2018;43:550–5.
[27] Purvis TE, Andreou E, Neuman BJ, Riley LH, Skolasky RL. Concurrent validity and responsiveness of PROMIS health domains among patients presenting for anterior cervical spine surgery. Spine 2017;42:E1357–65.
[28] Radcliff K, Davis R, Hisey M, Nunley P, Hoffman G, Jackson R, et al. Long-term evaluation of cervical disc arthroplasty with the Mobi-C© cervical disc: a randomized, prospective, multicenter clinical trial with seven-year follow-up. Int J Spine Surg 2017;11:31.
[29] Sasso RC, Smucker JD, Hacker RJ, Heller JG. Artificial disc versus fusion: a prospective, randomized study with 2-year follow-up on 99 patients. Spine 2007;32:2933–40.
[30] National Institutes of Health (NIH). PHYSICAL FUNCTION A brief guide to the PROMIS Physical Function instruments. Available at http://nihpromis.org/measures/domainframework2. Accessed April 15, 2019.
[31] National Institutes of Health (NIH). PAIN INTERFERENCE A brief guide to the PROMIS Pain Interference instruments. Available at http://nihpromis.org/measures/domainframework1. Accessed April 15, 2019.
[32] Riddle DL, Stratford PW. Use of generic versus region-specific functional status measures on patients with cervical spine disorders. Phys Ther 1998;78:951–63.
[33] Murphy HA, Warnick E, McEntee R, Nicholson K, Hollern DA, Stawicki C, et al. Which domains of the ODI best predict change in physical function in patients after surgery for degenerative lumbar spondylolisthesis? Spine 2018;43:805–12.
[34] Djurasovic M, Glassman SD, Dimar JR, Crawford CH, Bratcher KR, Carreon LY. Changes in the Oswestry Disability Index that predict improvement after lumbar fusion. J Neurosurg Spine 2012;17:486–90.
[35] Moses MJ, Tishelman JC, Stekas N, Jevotovsky DS, Vasquez-Montes D, Karia R, et al. Comparison of patient reported outcome measurement information system with neck disability index and visual analog scale in patients with neck pain. Spine 2019;44:E162–7.
[36] Sharma M, Ugiliweneza B, Beswick J, Boakye M. Concurrent validity and comparative responsiveness of PROMIS-SF versus legacy measures in the cervical and lumbar spine population: longitudinal analysis from baseline to postsurgery. World Neurosurg 2018;115:e664–75.